Bioinformatics, 37(8), 2021, 1176–1177
doi: 10.1093/bioinformatics/btaa781
Advance Access Publication Date: 10 November 2020
Applications Note

Structural bioinformatics

CHESPA/CHESCA-SPARKY: automated NMR data analysis plugins for SPARKY to map protein allostery

Hongzhao Shao[1], Stephen Boulton[2], Cristina Olivieri[3], Hebatallah Mohamed[1], Madoka Akimoto[1], Manu Veliparambil Subrahmanian[3], Gianluigi Veglia[3,4], John L. Markley [5], Giuseppe Melacini[1,2,*] and Woonghee Lee [5,6,*]

[1]Department of Chemistry and Chemical Biology, [2]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON L8S 4L8, Canada, [3]Department of Biochemistry, Molecular Biology and Biophysics, [4]Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA, [5]Biochemistry Department, National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, Madison, WI 53717, USA and [6]Department of Chemistry, University of Colorado Denver, Denver, CO 80217, USA

*To whom correspondence should be addressed.

Associate Editor: Ponty Yann

Received and revised on July 15, 2020; editorial decision on August 31, 2020; accepted on September 2, 2020

Abstract

Motivation: Correlated Nuclear Magnetic Resonance (NMR) chemical shift changes identified through the CHEmical Shift Projection Analysis (CHESPA) and CHEmical Shift Covariance Analysis (CHESCA) reveal pathways of allosteric transitions in biological macromolecules. To address the need for an automated platform that implements CHESPA and CHESCA and integrates them with other NMR analysis software packages, we introduce here integrated plugins for NMRFAM-SPARKY that implement the seamless detection and visualization of allosteric networks.

Availability and implementation: CHESCA-SPARKY and CHESPA-SPARKY are available in the latest version of NMRFAM-SPARKY from the National Magnetic Resonance Facility at Madison (http://pine.nmrfam.wisc.edu/download_packages.html), the NMRbox Project (https://nmrbox.org) and to subscribers to the SBGrid (https://sbgrid.org). The assigned spectra involved in this study and tutorial videos using this dataset are available at https://sites.google.com/view/chescachespa-sparky.

Contact: melacin@mcmaster.ca or woonghee.lee@ucdenver.edu

Supplementary information: Supplementary data are available at *Bioinformatics* Online.

1 Introduction

Protein allostery is one of the primary mechanisms for regulating biological function. It relies on long-range couplings between orthosteric and remote allosteric sites (Boulton and Melacini, 2016). Such couplings typically arise from transitions between macromolecular states with distinct structural and/or dynamic profiles. When allostery relies on subtle conformational and/or dynamical changes, the underlying allosteric networks often remain elusive to traditional structural determination methods. Protein NMR chemical shift perturbations offer a robust approach to mapping otherwise elusive allosteric networks. Specifically, variations in ^{1}H and ^{15}N resonance frequencies are exquisitely sensitive not only to nearest-neighbor binding contacts but also to subtle, yet functionally relevant, allosteric transitions underlying long-range effects. Both CHESPA and CHESCA assume that the dynamic conformational equilibria underlying the allosteric couplings occur in the fast chemical-exchange regime. In this case, the observed NMR chemical shifts are linear, population-weighted averages of conformer-specific values. Hence, it is possible to utilize peak positions to monitor shifts in local conformational equilibria.

The growing cohort of CHESPA and CHESCA applications (Akimoto *et al.*, 2013; Axe *et al.*, 2014; Chen *et al.*, 2017; Selvaratnam *et al.*, 2011, 2012; Shao *et al.*, 2020; VanSchouwen *et al.*, 2015; Walker *et al.*, 2019; Wang *et al.*, 2019; Xu *et al.*, 2019) illustrates the potential of these methods. To fully leverage this potential, both methods require a seamless integration with the PINE-SPARKY.2 and the I-PINE webserver (Lee and Markley, 2018; Lee *et al.*, 2019) for semi-automated protein chemical shift assignment and validation through 3D NMR triple-resonance experiments for the same protein in multiple bound or mutated forms. To enable such integration, we developed plugins and a webserver for the implementation and visualization of CHESPA and CHESCA through one of the most popular biomolecular NMR analysis suites: NMRFAM-SPARKY (Lee *et al.*, 2015) (Fig. 1 and Supplementary Fig. S1).

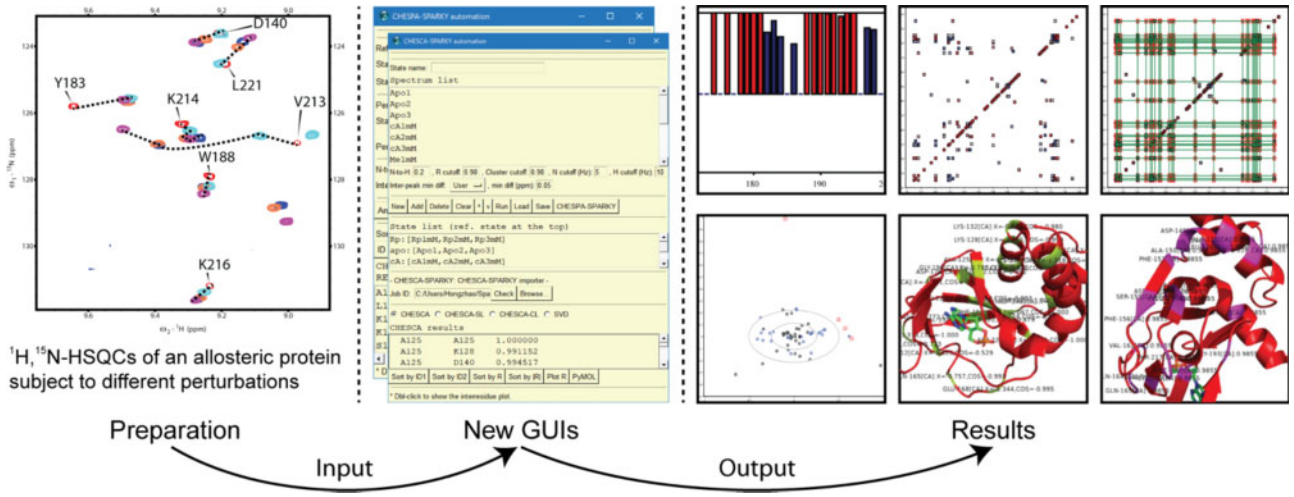

Fig. 1. General scheme for the CHESPA- and CHESCA-SPARKY plugins. The plugins start with NMR spectra of an allosteric protein system in multiple states. Then the assigned spectra are loaded into the new GUIs for the CHES(P/C)A plugins. Results are automatically computed and presented in the CHES(P/C)A plugins. Further details available in Supplementary Information

2 Implementation

The CHESPA (two-letter-code cH; Supplementary Fig. S2, left panel) and CHESCA (two-letter-code ch; Supplementary Fig. S2, right panel) plugins enable the semi-automatic visualization of allosteric networks starting from unassigned NMR spectra (Supplementary Methods). To execute CHESPA-SPARKY, both reference and perturbation vectors need to be defined as well as parameters, such as ^{15}N versus ^{1}H weights, between-state and inter-peak chemical shift cutoffs. CHESCA-SPARKY set up is similar but relies on a library of differentially perturbed protein samples, along with additional parameters, such as correlation coefficient cutoff values. Both methods are executed locally with agglomerative clustering and singular value decomposition run by the NMRFAM server. Results are plotted with the integrated NDP-PLOT module and can be highlighted through PyMOL molecular visualization if a PDB structure is provided.

3 Results and conclusions

Both CHESPA- and CHESCA-SPARKY plugins are illustrated through the analysis of a set of ^{1}H, ^{15}N HSQC NMR spectra of the RIα subunit 119–244 construct of the cAMP-dependent protein kinase (protein kinase A, PKA) and its four complexes with cAMP and analogs. The results of the demonstration are available in the Supplementary Figures S3–S8.

In conclusion, we have designed an integrated software platform that offers a seamless workflow from protein NMR spectra to allosteric network analysis and visualization. The new platform will enable the facile elucidation of allosteric mechanisms without having to export, compile and analyze multiple peak lists and chemical shift data through several nonintegrated scripts and software packages. Combined with the PINE-SPARKY.2 plugin, our CHES(P/C)A-SPARKY plugins significantly lowers the barrier for non-NMR specialists to map protein allosteric networks.

Acknowledgement

We thank Rashik Ahmed for helpful discussions.

Funding

This work was supported by the Canadian Institutes of Health Research Grant 389522 (to G.M.), the Natural Sciences and Engineering Research Council of Canada Grant RGPIN-2014-04514 (to G.M.), the National Institute of Health (P41GM103399 to J.L.M.; GM100310 and HL144130 to G.V.), the National Science Foundation (DBI-1902076 to W.L. and J.L.M.), and the University of Colorado Denver (STYPE 61193205 to W.L.).

Conflict of Interest: none declared.

References

Akimoto,M. *et al.* (2013) Signaling through dynamic linkers as revealed by PKA. *Proc. Natl. Acad. Sci. USA*, **110**, 14231–14236.
Axe,J.M. *et al.* (2014) Amino acid networks in a (β/α) 8 barrel enzyme change during catalytic turnover. *J. Am. Chem. Soc.*, **136**, 6818–6821.
Boulton,S. and Melacini,G. (2016) Advances in NMR methods to map allosteric sites: from models to translation. *Chem. Rev.*, **116**, 6267–6304.
Chen,H. *et al.* (2017) Elucidation of a four-site allosteric network in fibroblast growth factor receptor tyrosine kinases. *Elife*, **6**, e21137.
Lee,W. and Markley,J.L. *et al.* (2018) PINE-SPARKY.2 for automated NMR-based protein structure research. *Bioinformatics*, **34**, 1586–1588.
Lee,W. *et al.* (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, **31**, 1325–1327.
Lee,W. *et al.* (2019) I-PINE web server: an integrative probabilistic NMR assignment system for proteins. *J. Biomol. NMR*, **73**, 213–222.
Selvaratnam,R. *et al.* (2011) Mapping allostery through the covariance analysis of NMR chemical shifts. *Proc. Natl. Acad. Sci. USA*, **108**, 6133–6138.
Selvaratnam,R. *et al.* (2012) The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PLoS One*, **7**, e48707.
Shao,H. *et al.* (2020) Mechanism of action of an EPAC1-selective competitive partial agonist. *J. Med. Chem.*, **63**, 4762–4775.
VanSchouwen,B. *et al.* (2015) Mechanism of cAMP partial agonism in protein kinase G (PKG). *J. Biol. Chem.*, **290**, 28631–28641.
Walker,C. *et al.* (2019) Cushing's syndrome driver mutation disrupts protein kinase A allosteric network, altering both regulation and substrate specificity. *Sci. Adv.*, **5**, eaaw9298.
Wang,Y. *et al.* (2019) Globally correlated conformational entropy underlies positive and negative cooperativity in a kinase's enzymatic cycle. *Nat. Commun.*, **10**, 799.
Xu,Y. *et al.* (2019) Identifying coupled clusters of allostery participants through chemical shift perturbations. *Proc. Natl. Acad. Sci. USA*, **116**, 2078–2085.